Eur. J. Clin. Chem. Clin. Biochem.
Vol. 29, 1991, pp. 507–519

Laboratory Findings: Structure, Validity and Significance for Medical Cognitive Processes[1])

By *J. Büttner*

Institut für Klinische Chemie I im Zentrum Laboratoriumsmedizin, Medizinische Hochschule, Hannover, Germany

(Received April 15/May 21, 1991)

Summary: Modern medicine employs laboratory findings to a great extent in medical cognitive and decision processes. As supposedly "hard" data, the value of such findings is frequently incorrectly assessed. So far, no comprehensive general theory of laboratory findings has been available, although various subproblems have been dealt with. Firstly, the structure of a laboratory finding will be investigated in detail, proceeding from an analysis of the scientific language used for laboratory findings. The part played by laboratory findings in medical cognitive processes in making a diagnosis or prognosis will then be shown. Finally, attempts at characterizing the validity of a laboratory finding with the aid of statistical methods and information theory, as well as appropriate steps for checking the validity will be discussed.

Introduction

Laboratory findings are the product of work in the laboratory. In medicine, they are obtained daily in great numbers and employed by the physician in his practical activity at the sick-bed. The clinical chemist or laboratory physician compiling the laboratory findings is aware of the analytical errors which can falsify the findings; the physician at the sick-bed tends however to believe them to be "hard data". Both however consider laboratory findings to be empirical data essential in medical work. So far, not very much attention has been paid to the theoretical processes in the formation and evaluation of laboratory findings. In other words, laboratory findings have so far not been investigated from the standpoint of philosophy of science. Below, an attempt will be made at such an analysis. For this purpose, the laboratory finding must be seen in a larger context. The conceptual and methodic instrumentarium of clinical chemistry, biometry and informatics does not suffice; it must be amplified by drawing on other faculties, for instance analytical philosophy and semiotics (1).

[1]) Based on a lecture at the Heidelberg Colloquium of Medical Biometry, Informatics and Epidemiology, 21. 1. 1991.

Structure of a Laboratory Finding: Syntactics

In order to analyse the linguistic and notional structure of a laboratory finding, we shall proceed from a simple example. The essence of a laboratory finding is the description of an object property, in the simple example of table 1 a property of a specific urine. It

Tab. 1. Example of a laboratory finding

Ordered by: Dr. XYZ	Patient: Meyer, Paul
Sample taking: (date, time)	
Sample material: urine	
Examination wanted: test for glucose	
Finding: glucose in urine positive	

is characteristic of a laboratory finding to be the result of a scientific experiment, for instance a chemical analysis. In this paper, to simplify matters, we shall disregard the analytical chemical problems which are very important in practice and which can prevent exact acquisition of the object property sought. For setting forth the results of a laboratory investigation it will be expedient not to use colloquial language but

to employ a more highly formalised *"scientific language"* (2). The sentences in which this is done are referred to as *"observation sentences"* or *"protocol sentences"*, for instance: *"The specimen x of patient y had the property z at the time t"*. Such a sentence is the mental reflection of a fact, in our case an object property. The essence of the statement can be expressed in the language of mathematical logic in sample manner by a predicate variable G and an individual variable u. If the predicate variable G means *"contains glucose"* and the individual variable u means a specific urine, then we can express the observation *"urine u contains glucose"* as $G\ u$. However, the relationships are usually more complex, in that the original observation is "transformed" to the final result by one or more conclusions. In the simple example of urine examination for glucose, the immediate observation is a change in colour. On the basis of existing knowledge of the analysis method applied, it is possible to derive from this the statement that the urine contains glucose. For this derivation of one sentence from another, a transformation rule is necessary. The procedure which can be adopted here is the scheme of the *"modus ponens"* of classical logic (3). This requires a general regularity, for example a "if-then implication", expressing general knowledge on the analysis method employed. In the case of our simple example this implication could read: *"it is true for all individuals that: if colouring occurs then glucose is present"*. Or, using the notation of mathematical logic

$$\wedge x\ [C\ x \Rightarrow G\ x]$$

The symbol $\wedge x$ denotes the all-operator ("for all x ... is true"). The predicates denote: C "colouring", G "glucose". The individual variable x means *"analysis sample"*.

From this regularity the conclusion $G\ u$ can be derived, which is the actual statement of the laboratory finding according to the following scheme:

Premises	$\wedge x\ [C\ x \Rightarrow G\ x]$	*Regularity of the chemical analysis*
	$C\ u$	*Initial conditions*
Conclusion	$G\ u$	

(The individual variable u stands for a specific urine).

With complicated analysis methods the transformation rule often contains several implications in succession. The procedure for deriving the transformation rule corresponds to the general scheme of a scientific explanation which was investigated in particular by *Hempel* and *Oppenheim* (4, 5) (see below for more details).

The efficiency of a language in representing object properties increases if adjectives (for example "red", "cloudy") or numbers are used. In a scientific language of the type we use for laboratory findings, by *metrisation* (tab. 2) a greater precision of the scientific statement can be achieved (6). Metrisation does not mean that the object property is defined differently;

Tab. 2. Metrisation

	Classification	Comparison	Quantification
Basic operation	Determination of identity	Description using relations "smaller than", "larger than", "equal to"	Determination of equality of a ratio or an interval
Scale	Nominal scale	Ordinal scale	Interval scale, ratio scale
Example	*Substance present or not present*	*Mass of substance A equal, smaller or larger than mass of substance B*	*Amount of substance in relation to standard*

all that is changed is our system of concepts, i. e. the language in which we express our finding. In the simplest case, the object properties are *classified* on the basis of specific qualitative characteristics. The fundamental operation is the determination of the identity of the object property with the property of a defined class. Where possible, a simple comparison method for the respective object property can be developed which permits the determination of relationships, for example "greater or smaller than". This *comparison* is the next step of metrisation. The highest degree of metrisation is achieved by devising a quantitative method, i. e. by a measurement in the true sense. This is called *quantification*. Thereby it is possible to assign numbers to the object property. This is frequently done by determining a ratio, for example the measurement result for the object to be examined related to that of an exactly measured reference material.

The introduction of metrical concepts requires exact definition of a quantity for the object property and the creation of suitable scales (7, 8). In table 3 the definition of a quantity in clinical chemistry will be explained. By international agreement, a quantity is represented by the elements *"system"* (investigated material), *"analyte"* (component to be determined), *"kind of quantity"* (e. g. substance concentration) and

Tab. 3. Definition of a clinical chemical quantity

A clinical chemical quantity consists of the following elements:

System	*Material to be examined*
Analyte Kind of quantity	*Component and property to be measured*
Unit	*Scale*

Example:

Serum — potassium, amount of substance concentration, 4.6 mmol/l

"unit" (e. g. mmol/l). The choice of the kind of quantity governs the nature of the scale and the unit defines the scale exactly (9, 10). For a typical example of the problems which may arise in setting up a metrical system for a chemical analysis, a recent paper on measurement of biological substances by means of immunoassays (11) should be consulted.

Usually, laboratory findings as a description of an object property are not stationary quantities independent of time. For this reason, a general conceptual definition of the laboratory finding must also contain the dimension time. Theoretical models to describe the time dependence of laboratory findings can be very complex. Attention will be drawn here only to the simple possibility of describing the time profile of laboratory findings in a pathological process as a *Markov* chain (12). The typical course of a disease, e. g. of a certain carcinoma, can be described by a number of states or stages I, II, III ... The transition from one stage to another, e. g. I → II, may be characterized by the conditional probability P (I | II). The *Markov* graph contains all of the transitions and their probabilities. In figure 1 the stages of the prostatic cancer are defined by the outcome of enzyme determinations, e. g. *"alkaline phosphatase non-pathological"* and *"acid phosphatase pathological"*. The numbers denote the conditional probabilities (taken from l. c. (13)).

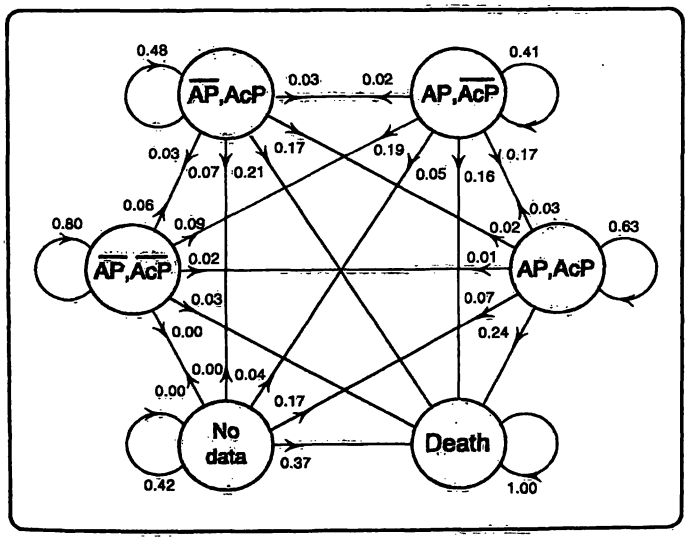

Fig. 1. *Markov* model of prostatic cancer.
Explanation: $\overline{AP}$ = alkaline phosphatase not elevated, AP = alkaline phosphatase elevated, $\overline{AcP}$ = acid phosphatase not elevated, AcP = acid phosphatase elevated.
Numbers denote transition probabilities.
From *Meyers* et al. (13).

It is apparent from this short consideration of the structure of laboratory findings that the latter represent observation sentences obtained from an experiment. Various rules must be followed in order to represent correctly these observation sentences and the propositions derived from them. It must be ensured that the laboratory findings are "syntactically well formed observation sentences".

Significance of a Laboratory Finding: Semantics

In the analysis of the linguistic and theoretical structure of laboratory findings, the "content" or the "meaning" of a laboratory finding have so far been disregarded. The physician at the sick-bed wishing to employ the laboratory findings is not interested in their structure. He wants to know their *"meaning"*. The "meaning" of a laboratory finding can be different depending on the medical use. Therefore it appears advisable to take a short look at the medical uses made of laboratory findings before investigating the term *"meaning"* (tab. 4).

Tab. 4. Medical use of laboratory findings

Diagostic use
- Classification of disease
- Determination of etiology and patho-mechanism
- Examination of the state of the patient
- Searching for risk factors

Prognostic use
- Prognosis with respect to exit (death, cure)
- Prognosis with respect to course
- Prognosis with respect to risk of therapy
- Prognosis with respect to future diseases

Use in connection with therapeutic measures
- Selection and control of efficiency of therapeutic measures

What is the *"meaning"* of a laboratory finding? How in fact does the finding obtain its meaning at all? The instruments of modern semiotics, the theory of signs, are a great help in answering these questions (for more information about semiotics see for example: o. c. (14—19)). Just like clinical symptoms, laboratory

findings are *signs*, i.e. *phenomena standing for something else*. These signs stand for a disease; they indicate a disease.

In semiotics a distinction is made (20) between the *sign*, i.e. the phenomenon observed, and the *significatum*, i.e. what the sign indicates or what it "signifies". The significatum is a concept, that is a mental construct, existing only in the interpreting mind of the physician. In many cases the significatum relates to a concrete object. For this, the term *denotatum* is used. The sign, significatum and denotatum are often expressed in their mutual relationships in the form of the *"Baldinger Triangle"* (21) which is shown in figure 2, using the example of a "diagnostic sign".

Now, how is a meaning allocated to a sign? This operation is called "process of signification" or *"semiosis"*. *Charles Morris* (22), one of the founders of semiotics, distinguished between three subareas of semiotics and introduced for them the terms *"syntactics"*, *"semantics"* and *"pragmatics"*. Syntactics investigate the structure and mutual relationship of signs, i.e. the topic discussed above. Semantics on the other hand relate to the relationship of a sign and its significatum or denotatum. Finally, pragmatics investigate the use of the sign by the user.

Below, the process of signification for a laboratory finding will be set forth in somewhat more detail. We take as an example a clinical chemical finding which is used for diagnostic purposes. The process of signification takes place in stages and a distinction can be made between three levels; these will be denoted *"technical level"*, *"biological level"* and *"nosological level"* (fig. 3).

As explained above, on the technical level the "syntactically well formed" observation sentence describing an object or material property is developed.

On the biological level, different classifications take place which are referred to in clinical chemistry as "longitudinal evaluation" and "transversal evaluation" (23, 24). The former investigates the relationship of an actual finding to the preceding findings for the same patient, i.e. the change with time. In the transversal evaluation, on the other hand, the actual finding is compared with a reference population. This gives the individual laboratory finding the meaning of *"pathological"* or *"non-pathological"*. The reference population, for example a population of clinically healthy persons, must be defined by suitable external criteria. The transversal evaluation takes place by means of a reference interval which is defined as the central 0.95 fraction of the values of the reference population (25).

On the third level, the nosological level, the sign is allocated to the significatum, i.e. the disease. Here, an interesting change has taken place in the course of the history of medicine (for a discussion of the historical development of "chemical signs" in medicine see l. c. (26)). Originally, a series of empirically obtained signs was simply given the name of a disease. The sign was used without any causal explanation as "indication" of a disease. This method, first used by *Hippocrates*, could be called the "syntactic" method (27); the sign pattern is allocated to a disease name. Modern attempts at computer-aided diagnosis frequently employ the same principle. With increasing knowledge, a start has been made at explaining the signs observed on the basis of a theory of the disease. This is the "semantic" method, the beginnings of which can be traced back to *Galen* but which was not fully developed until the scientific medicine of the 19th century. The semantic method requires the definition of a *"morbus"*. In scientific medicine, *morbus*

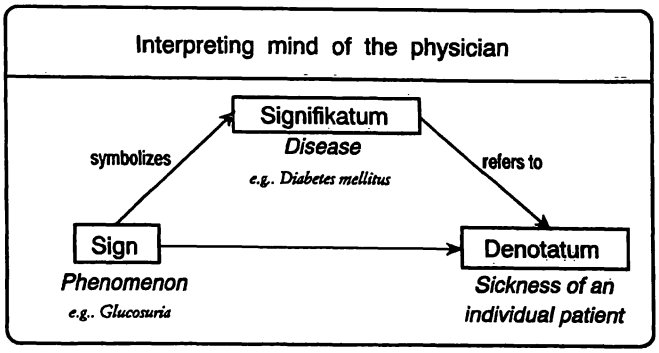

Fig. 2. Laboratory finding as a medical sign.

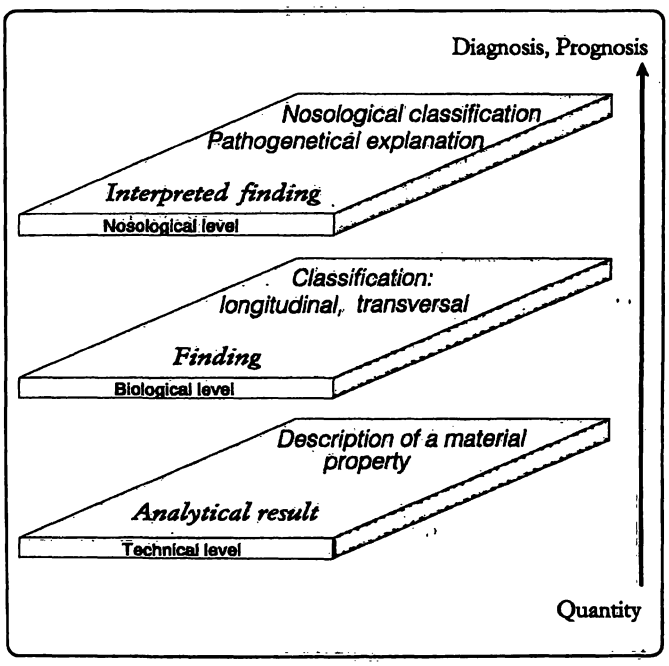

Fig. 3. Signification process for a laboratory finding.

is the term used when etiology and pathogenesis are known and uniform. If either the etiology or the pathogenesis or both are not clear or not uniform, the term *"syndrome"* is used (28). The allocation of a sign requires the "causal" explanation of the sign based on the pathogenesis of the morbus. It is only a sign interpreted in this manner which can be regarded as of full diagnostic value. Now, what is the "causal" explanation of a laboratory finding? A simple example will be taken: The leading laboratory findings in *Galactosaemia* are the elevated concentration of galactose in blood and the excretion of galactose in urine (galactosuria). This is explained by the genetically induced absence of the enzyme, galactose-1-phosphate uridyl transferase, which can be demonstrated in erythrocytes. This type of *"inborn error of metabolism"* (*A. E. Garrod*) with a monogenetic defect shows a very clear relation between *"cause"* and *"sign"*.

An important step in the process of signification for a diagnostically used laboratory finding is therefore the explanation of the finding on the basis of the pathogenesis of the morbus. What however does "explanation" mean? The cognitive process on which a scientific explanation is based may be explicated as follows (29—31): a process or a phenomenon is to be explained which is described by an observation statement. For the explanation, *universally valid laws L* and *antecedence conditions* (specific marginal conditions) *A* are required. These two together form the so-called *explanans*.

Explanans	universally valid laws	$L_1, L_2, L_3 \ldots$
	antecedence conditions	$A_1, A_2, A_3 \ldots$
Explanandum	proposition	E

The explanation consists in deductively deriving a statement or proposition as explanandum from the explanans. This is the so called *deductive-nomological scheme (DN scheme) of the explanation developed by Hempel & Oppenheim* (32).

The example discussed above may be represented as follows using this scheme:

Example: Galactosaemia

L: $\wedge x\,[\neg T\,x \Rightarrow H\,x \wedge U\,x]$

A: $\neg T\,a$

E: $H\,a \wedge U\,a$

Explanation:
T = Galactose-1-phosphate uridyl transferase present
H = hypergalactosaemia
U = galactosuria
a = patient y
$\wedge x$ = "all-operator"
$\wedge$ = conjunction "and"
$\neg$ = "not"

In medicine, universal laws are rather the exception. This is due to the great complexity of biological systems and the methodical uncertainty resulting from that. In addition, as we know today, certain laws in the molecular range are of an irreducible statistical nature. Pathophysiological or pathobiochemical maxims which we use for explanations are therefore frequently only statements of probability. A simple example will make this clear: It is known from empirical investigations that with patients displaying an increase in the glucose concentration in the blood after standardized administration of glucose, the metabolism of glucose is *frequently but not always* impaired. In a specific case a causal relationship may be assumed but not proved with certainty, because the system of homoeostatic regulation of glucose is very complex.

In such a case the previously discussed DN scheme of the explanation based on classical logic cannot be applied. Instead, *Hempel & Oppenheim* have proposed the scheme of *"inductive-statistical explanation"* (IS explanation) (33).

Premises	statistical "laws"	$S_1, S_2, S_3 \ldots$
	antecedence conditions	$A_1, A_2, A_3 \ldots$
	support	
Conclusion	proposition	E

Because of various epistemological difficulties presented by this concept, *Stegmüller* (34) prefers the weaker term *"inductive-statistical substantiation"* ("Begründung"). The decisive point is that a universal law is replaced by a probability statement. It is then however no longer possible to deduce logically an explanandum as in the DN explanation. Instead of this, the *"degree of confirmation"* is given with which a proposition is supported by the available knowledge contained in the premise. As is known, *Carnap* (35) expressed this "degree of confirmation" as *"inductive probability"* or *"probability$_2$"*. In contrast to the DN explanation, the substantiation by the IS scheme requires the exact specification of the statistical "laws" employed. This means that the knowledge basis on which these "laws" are based must be exactly defined. Also, the conclusion of an IS substantiation cannot be formulated as a modal statement (e. g. "very probable", "less probable", etc.), because these expressions say something about the relationship of the premises

to the conclusion but are not themselves part of the conclusion. These restrictions can result in considerable problems in the practical application, especially in the development and use of expert systems.

Using this scheme, we get for the example mentioned above

Example: glucose tolerance
S: $P(A|G) = q$
A: $G\ a$

---

q

E: $A\ a$

Explanation: Degree of confirmation
= probability q
A = assimilation of glucose impaired
G = glucose concentration in blood elevated after standardized administration of glucose
a = patient y

As a result of our explication of the meaning of a laboratory finding we can state: The laboratory finding obtains its meaning, i.e. its significatum as sign of a disease, through a complicated multistage process of signification. The aim is the explanation on the basis of the pathogenesis of the morbus concerned. Explanation in the strict sense is a deductive process. In medicine, it is frequently replaced by the weaker inductive-statistical substantiation.

The Diagnostic Process

Just like other medical signs and symptoms, laboratory findings are processed at the sick-bed in special medical cognitive processes. Such cognitive processes are for instance the compiling of a diagnosis or postulation of a prognosis. Diagnosis and prognosis are in turn again the starting point or *"indication"* for action (36) on the part of the physician (for the theory of diagnosis see for example: o.c., l.c. (37–42)).

The basic process of making a diagnosis may be described as a classification. On the basis of a multidimensional finding vector the disease of a patient is classified as a certain morbus (or as a certain syndrome). Corresponding theoretical models have been developed in particular in conjunction with "computer-aided-diagnosis". It would be going too far to discuss this in detail here (for literature summary see l.c. (43)).

The cognitive processes in making a diagnosis are very complex. *Hartmann* once said: *"Medical thinking is as a rule a search process, attempt, dismissal, assumption and confirmation"* (44). The diagnostic process thus cannot be fully described by a simple model. In addition, there is the uncertainty about the concept of diagnosis itself. The classical diagnosis concept of scientific medicine is relativated in its significance by the current discussion on the philosophy of science. Also, today in practical medicine the diagnosis process is frequently only carried on until clear therapeutical alternatives become apparent.

In our context, the question which interests us most is at what point of the diagnostic process laboratory findings can be used. This can be illustrated very clearly using a two-stage scheme developed by *Medawar* (45, 46) and referred to as a *"hypothetico-deductive scheme"* (fig. 4). In the first phase, which we shall refer to as *"hypothesis formation"*, a tentative diagnosis is made. For this phase, which generally takes place in stages, no strict logical rules apply. The physician attempts to combine the symptoms observed and the findings of the concrete case in an assumed or tentative diagnosis. In this phase, laboratory findings provide more or less concrete indications. The hypothesis formation is carried on against the background of the clinical experience of the physician. Of great importance is the subjective or objective probability assumed for the occurrence of the suspected disease. For example the tentative diagnosis "virus influenza" will be influenced by current information on an influenza epidemy. From the semiotic point of view, in forming the hypothesis the significatum, i.e. the disease, is deduced from the sign. The sign is used "semiotically".

The second phase is the *confirmation of the hypothesis* formed. For this phase, the strict rules of logical deduction apply. The confirmation of the tentative diagnosis may be formulated as explanation.

Explanans	universally valid laws symptoms, sings and findings in the clinical picture	$L_1, L_2, L_3 \ldots$
	antecedens conditions diagnosis for patient y	$A_1, A_2, A_3 \ldots$
Explanandum	symptoms, signs and findings for the patient y	E

Example (according to *Wieland* (47)):
L: If acute glomerulonephritis, then haematuria, oedema, hypertension, proteinuria
A: patient y has acute glomerulonephritis

E: Patient y exhibits haematuria, oedema, hypertension, proteinuria

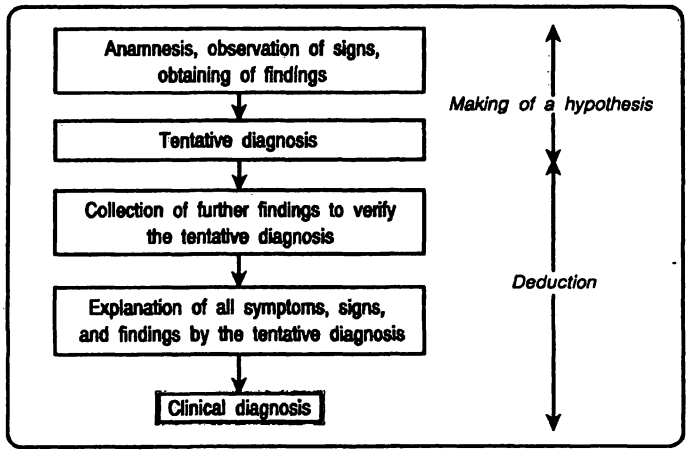

Fig. 4. Diagnostic process as hypothetico-deductive scheme.

However, as *Wieland* has pointed out (48), it must be remembered that medical diagnoses, in contrast to scientific explanations, are *singular statements* which relate only to a specific patient. Thus, the symptoms and findings of this patient are explained from the knowledge of the morbus and the tentative diagnosis made in the first phase. In other words, the concept of the disease is assigned to the illness of this patient. By the way, the restrictions already discussed apply to the deductive phase of the diagnostic process if statistical "laws" are employed in the explanans, as is the rule in medicine. From the semiotic point of view, in the deductive phase a conclusion is drawn from the designatum, the disease of a certain patient, to the sign; the sign is employed "nosologically".

This brief outline of the diagnostic process will have to suffice here. This is not the place to go into details, in particular the important problem of differential diagnosis.

The Prognostic Process

Let us now briefly consider the making of a prognosis, which is based on a very different question compared with the diagnosis (49). The physician attempts to make a prediction on the outcome or course of a disease for a certain patient. The prognosis, which in medicine in previous centuries attained a greater significance than the diagnosis, has only recently received increased attention again, for example in chronic diseases or in intensive care medicine, where prognostic conclusions regarding the acute danger are an urgent necessity without previously making a diagnosis. Decision on high-risk therapeutical measures (e. g. transplants) also requires prognostic statements, either to make decisions or to evaluate steps taken.

From the logical point of view, today a prognosis is dealt with in the same manner as the explanation already repeatedly mentioned: from laws and actual antecedence conditions (here also referred to as *"starting point"*) the *"end point"* of the prognosis is deductively concluded.

Explanans	universally valid "laws"	$L_1, L_2, L_3 \ldots$
	antecedens conditions "starting points of the prognosis"	$A_1, A_2, A_3 \ldots$
Explanandum	prognosis "end point of the prognosis"	P

In a manner very similar to the situation for a diagnosis, the "laws" available for prognoses are of a mainly statistical nature. Further characteristic of prognoses is that frequently, although not always, the dimension of time is contained in the laws used.

A simple example will be considered (according to *Pui* et al. (50)): In children with acute lymphatic leukaemia the level of serum lactate dehydrogenase (S-LDH) will be elevated before treatment if the survival time is short. This can be formulated as a "statistical law". Now we consider a case of a child *y* with acute lymphatic leukaemia having elevated serum lactate dehydrogenase. From the two premises the "end point" of the prognosis is derived and in our example this is for instance a statement how many years the patient will survive:

L: Survival time dependent on S-LDH-activity before treatment
A: Actual activity of S-LDH in patient y

P: Probability statement on survival time of patient y

Validity of Laboratory Findings

Laboratory findings are the result of a scientific experiment. In hospitals and clinics this has given them the reputation of "hard data". A more exact examination shows, however, that it is possible for laboratory findings to embody considerable uncertainties. These uncertainties influence the medical cognitive processes discussed and lead to incorrect conclusions. In practice, this raises the question of the validity of laboratory findings (51–53). Here, a *"valid laboratory finding"* will be considered quite generally to be *a result which correctly answers the question put* by the physician.

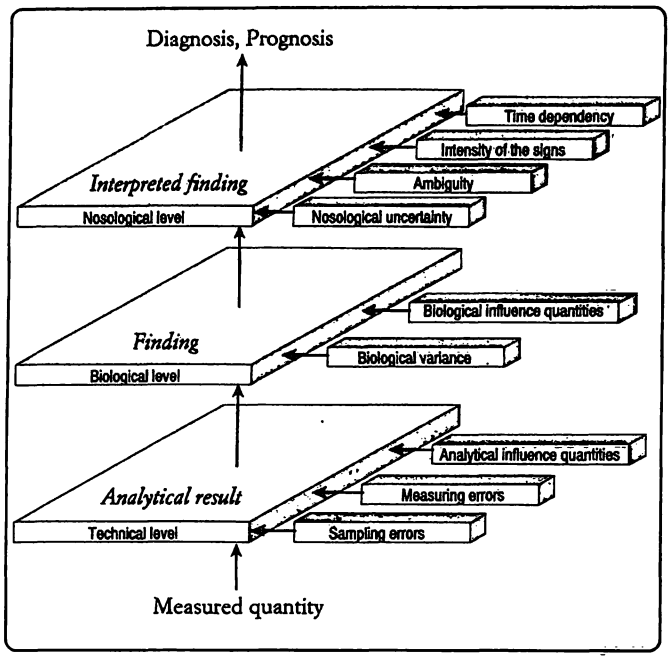

Fig. 5. Uncertainty of laboratory findings.

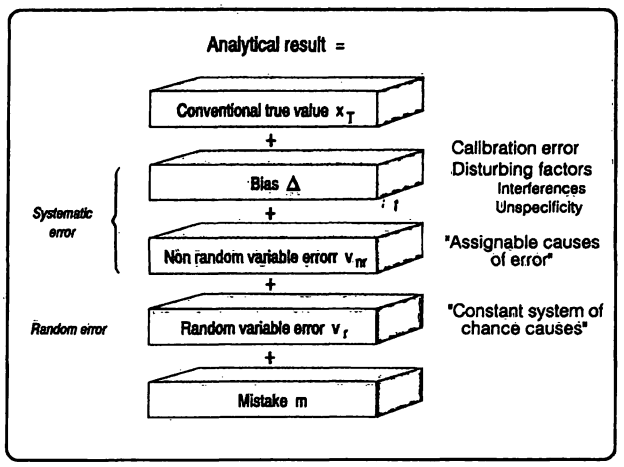

Fig. 6. Analytical error of clinical chemical analyses.

To analyse the validity of laboratory findings (fig. 5) the three-stage scheme already employed when discussing the process of signification (cf. fig. 3) will be taken as a basis. On the technical level the *analysis result* is obtained. This is where the errors due to the analytical methods occur. On the biological level, the *finding* is derived from the analysis result. Here, in particular the biological variance of the quantity investigated leads to uncertainties. Finally, on the nosological level, where the *interpreted finding* is formed, uncertainties arise from the definition of the morbus and the inadequate or incorrect pathophysiological explanation; however, uncertainties also arise due to the different intensity of the signs and their ambiguity. Thus, a valid laboratory finding must

(1) correctly describe the property intended to be measured,
(2) be correctly assessed on the biological level, and
(3) properly assigned to a morbus or syndrome.

This great variety of influences and uncertainties makes it necessary to develop measures suitable for the checking and assurance of the validity of the laboratory findings on all levels of the signification process. These may be measures for generally testing ("evaluating") the examination method and monitoring it. However, methods for validating the individual finding of a specific patient are also required.

Errors which occur in compiling laboratory findings on the technical level, the "analysis errors", can be described very well by an additive model (fig. 6) (54). To monitor the errors on the technical level, sophisticated statistical control methods have been developed which are employed generally today as *"statistical quality control"* (55–57). It is more difficult to detect the errors in the preanalytic phase, i.e. during the preparation before the actual analysis (58). Examples are the errors that occur when taking a blood sample or due to incorrect storage of the sample. One possibility of detecting such errors is provided by the methods of *"plausibility control"* (59, 60). Here, the individual laboratory finding is checked by comparison with a large data base.

The cause of uncertainties of laboratory findings on the *biological level* is biological variation. Their determination and the estimation of their influence on the validity of a laboratory finding is a classical problem in biometry. The uncertainties due to biological variation have considerable effects on the longitudinal evaluation and transversal evaluation, those two operations in compiling a laboratory finding which we have already discussed in connection with the process of signification.

What possibilities do we have for improving the validity of laboratory findings on the biological level in view of the fact of biological variation? By statistical analysis of suitable data material, it is possible to identify for a quantity a number of factors which are referred to as *"biological influence quantities"* (61–65).

The variation of laboratory findings may be described by a simple additive model (66):

Total variance	= intra-individual variance	+ inter-individual variance	+ analytical variance	+ residual variance
σ^2	$= \sigma^2_{BI}$	$+ \sigma^2_{BP}$	$+ \sigma^2_A$	$+ \sigma^2_{res}$

The intra-individual variance σ^2_{BI} and the inter-individual variance σ^2_{BP} are decisively governed by the biological influence quantities. This may for example be a genetically induced influence or physical activity or the intake of certain substances with food, such as fat or toxic substances in tobacco, alcohol, coffee etc. Some of the biological influence quantities depend on time, such as aging or certain biorhythms. The aim is to detect all these influence quantities and take account of them when examining a patient, standardizing them whenever possible. Thus, the old rule of the clinician of taking blood samples for clinical chemical investigations in the morning from a fasting patient leads to such a standardization. The concept of the biological influence quantity also has consequences regarding the correct determination of reference values as required for transversal evaluation. It is essential to select the reference persons carefully and standardize any possible biological influence quantities in the reference population using the same protocol for testing, which is then applied in the examination of clinical patients (67).

The greatest uncertainties in laboratory findings are on the nosological level. We shall discuss this by taking as example the interpretation of a laboratory finding used for diagnostic purposes. Here, many causes can be found which result to a particular extent in uncertainties.

The *"nosological uncertainty"* arises in the process of signification of the laboratory finding when the morbus or syndrome, which are of course mental constructs, are not clearly defined and delineated with respect to similar diseases. However, uncertainties arise for the laboratory finding particularly when the causal explanation of the finding based on the pathogenesis of the morbus has not been unequivocally proved experimentally.

Another source of uncertainty is the *ambiguity of the findings*. The ideal diagnostic sign is *"pathognomonic"*, i.e. it occurs only in a very definite disease. Unfortunately, such ideal signs are rather the exception. In clinical chemistry we find them for instance in genetically induced metabolic disturbances, such as the phenylpyruvic oligophrenia: the finding of phenylpyruvic acid passed in the urine occurs only in this disease and is proof of the disease. However, most findings are equivocal and can be interpreted in various ways.

The *different intensity of the findings* in one and the same disease in particular leads to uncertainties. For a clinical chemical finding this may for example be due to a "pathological" substance, such as a tumour marker, being produced in different quantities and consequently detectable in the blood in fluctuating concentration.

Often related to this is the fourth cause of uncertainty, the *dependence of a finding on time*, for example during the course of a disease. In clinical practice attempts are made to reduce this uncertainty by defining "stages" within the course of the disease.

The uncertainties on the nosological level outlined can greatly mislead the physician at the sick-bed in his interpretation of laboratory findings. It is therefore absolutely essential to carefully *validate* laboratory investigations before their clinical introduction. Apart from this general examination of the *methods*, an examination and evaluation of the laboratory *findings* of a specific patient is of course essential in every case.

How can clinical laboratory examinations or diagnostic examination methods in general be assessed as regards their diagnostic value? The reason for working out suitable validation methods was the broad introduction of diagnostic methods for "screening" large groups of the population in the period following the Second World War, because here expensive classification errors made themselves particularly noticeable.

The basic problem, which we shall refer to as the *"diagnostic test problem"*, is illustrated in figure 7. A "positive" or "negative" test result is to be used to make the classification "sick" (or "not healthy") or "not sick" (or "healthy"). The aim is to find suitable quantities or "figures of merit" expressing the efficiency of this classification.

It will not be possible to deal systematically here with the large number of validation methods which have been described for diagnostic examinations. Attention is drawn to some general studies (68–73), in which

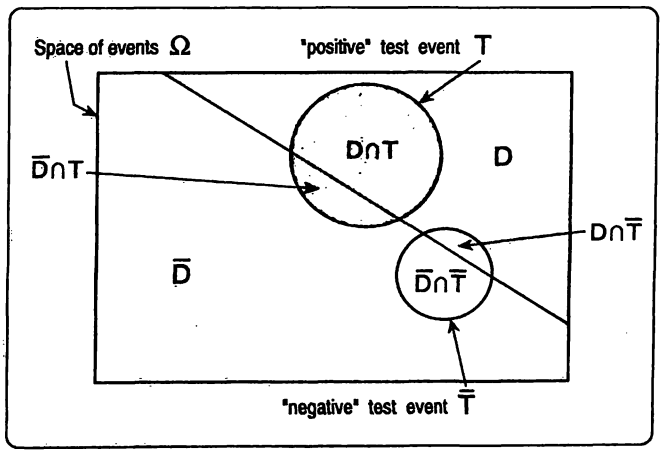

Fig. 7. Diagnostic testing problem.
Explanation: D = diseased, D̄ = not diseased, T = test result positive, T̄ = test result negative.

these methods are critically explained. Here, three such validation methods will be briefly considered exemplarily.

The classic method originates from a paper by *Yerushalmy* in 1947 (74), and is explained in figure 8. It is the simple case of a qualitative examination with dichotomic outcome ("pathological", "not pathological"). Examples are the detection methods for protein or glucose in urine. To characterize the validity the terms *"diagnostic sensitivity"* and *"diagnostic specificity"* are introduced which represent conditional probabilities. The great practical advantage of this concept is that it allows the physician to choose an examination method corresponding to the question to be answered. In the phase of hypothesis formation when making a diagnosis, he will choose as "search test" a sensitive test, but in the confirmation phase he will choose instead a specific test with high proof value as the "confirmation test".

A completely different approach for solving the validation problem originates from the signal detection theory which tackles the detection of a signal in the presence of noise (75, 76). The essential instrument is the *"receiver operating characteristic curve"* (ROC), which represents the frequency of the "hits", i.e. the "true positive" results, in dependence upon the frequency of the "false alarms", i.e. the "falsely positive" results (fig. 9). A practical advantage is to be seen in the possibility of defining the optimum setting of the tests, which can be achieved by shifting the decision limits.

A third approach proceeds from the consideration that the diagnostic examination provides the physician with information. Information is defined in the

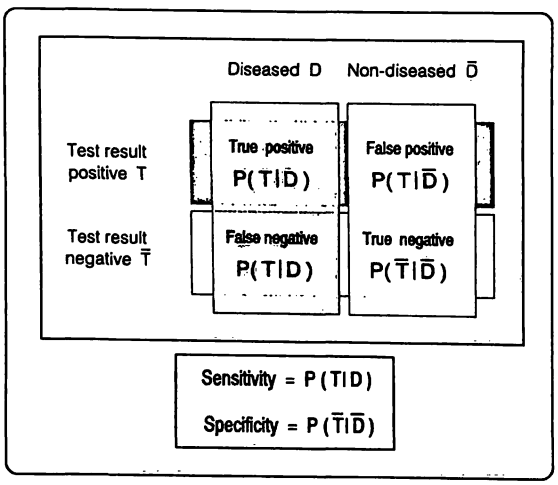

Fig. 8. Validation of a dichotomic test.
Explanation: D = diseased, D̄ = not diseased, T = test result positive, T̄ = test result negative.

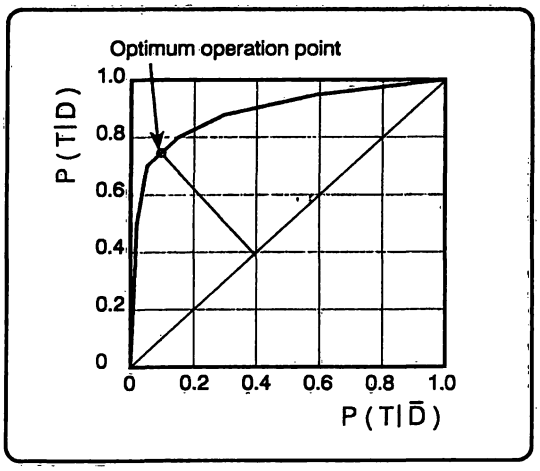

Fig. 9. Receiver operating characteristic.
Explanation: D = diseased, D̄ = not diseased, T = test result positive.
Conditional probabilities are denoted as P(X | Y) (probability of X given Y).

theory of information (77–79) as the elimination of uncertainty, for example by reception of a message. Messages on rare events have a higher information content than those on frequent events. The *information entropy*, a logarithmic function of the event probabilities, furnishes a quantitative measure of the information. A diagnostic test may now be represented as a model in the form of a communication channel (fig. 10). The information gained by the test, the so-called *transinformation I*, can be calculated as entropy difference and used as a measure of quality for the test (fig. 11). The advantage of the information theory approach lies in the extensive independence from any assumption on the nature and distribution of the data.

That completes our consideration of these examples, intended to indicate ways of validating diagnostic investigations on the nosological level. It must however be pointed out that the validation methods available at present cannot be applied without careful study to all investigation methods. Particular problems are presented by quantitative investigations. The procedure usually adopted there is to classify the results using a decision limit, thus transforming the quantitative test to a qualitative test with dichotomic outcome.

Finally, in table 5 the methods discussed for validating laboratory examinations and laboratory findings will be summarized once again. A distinction is made here between overall measures, with which the examination methods can be checked in general, and specific measures which can be applied to validate the individual laboratory findings.

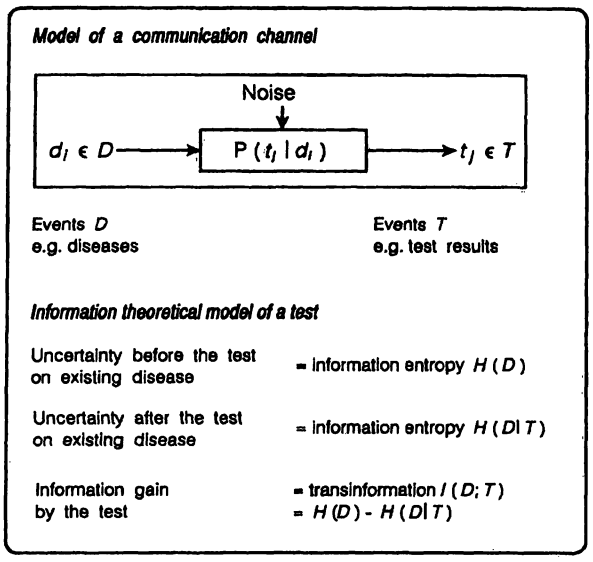

Fig. 10. Information theoretical model of a test.
Explanation: Entropies $H(X)$ and conditional entropies $H(X|Y)$ are given in a notation similar to the probabilities (see fig. 8).

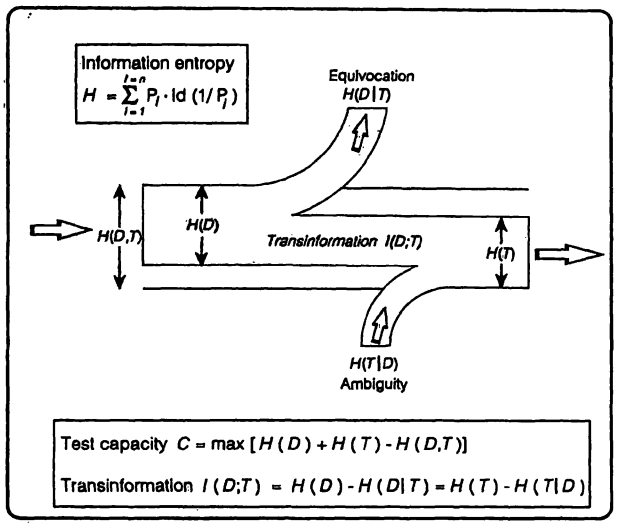

Fig. 11. Transinformation of a test (*Berger* diagram).
Explanation: Entropies $H(X)$ and conditional entropies $H(X|Y)$ are given in a notation similar to the probabilities (see fig. 8).

Conclusion

This investigation was undertaken with the objective of developing and precisely defining the concept of medical laboratory findings. In medicine, usually only vague ideas exist about the structure, formation and mental processing of laboratory findings. This made an exact analysis from the point of view of the philosophy of science and semiotics appear desirable. Such an analysis may help the clinician and the clinical chemist to better understanding of the theoretical basis of laboratory findings and their validity. Moreover it could be the basis for the development of effective expert systems. Any diagnostic system based on "artificial intelligence" must reconstruct the cognitive processes which result in a diagnosis. The structure of the scientific language used to describe the laboratory finding and the logic deductions should be reproduced correctly by the expert system.

Our investigation has shown that laboratory findings are observation sentences which describe an object property. They are expressed in a relatively highly formalised and metrised scientific language. Laboratory findings obtain their "meaning" in a multistage process of signification, which has been analysed in detail. Through this process laboratory findings become signs symbolizing or indicating a significatum, for example a morbus. The use of these signs was studied by way of example with reference to the processes of making a diagnosis and prognosis.

Finally, the numerous possibilities of uncertainties in these signs arising during the process of singification were considered. Some methods of validating and checking have been outlined.

In closing, it is pointed out that the theoretical analysis dealt with here was restricted to the example of the laboratory finding made by an experimental examination using methods of analytical chemistry. The structure of such laboratory findings is relatively simple. Further investigation is necessary to determine whether the results of other medical investigation methods may be dealt with in a similar way.

Tab. 5. Validation of laboratory findings

	Measures for validation	
	General	*Individual case*
Technical level *Quantity* ↓ *Analytical result*	Correct definition of the quantity to be measured Evaluation of the analytical method	Quality control
Biological level *Analytical result* ↓ *Finding*	Determination of correct reference intervals	Plausibility control Longitudinal evaluation Transversal evaluation
Nosological level *Finding* ↓ *Interpreted finding*	Evaluation of diagnostic and prognostic efficiency Pathophysiological explanation	Estimation of predictive value Evaluation with respect to the whole clinical picture

References

1. Büttner, J. (1991) Semiotik diagnostischer und prognostischer Untersuchungen. In: *Künstliche Intelligenz. Symposium anläßlich des 65. Geburtstages von Prof. Dr. Dr. H. Keller.* GIT Verlag GmbH, Darmstadt, pp. 45–59.
2. Carnap, R. 81934/1968) *Logische Syntax der Sprache.* 2nd Edition. Springer-Verlag, Wien, New York.
3. See textbooks of modern logic, e. g.
   (a) Tarski, A. (1977) *Einführung in die mathematische Logik.* 5th Edition. Vandenhoeck & Ruprecht, Göttingen.
   (b) Bucher, T. (1987) *Einführung in die angewandte Logik* [Sammlung Göschen 2231]. Walter de Gruyter, Berlin, New York.
   (c) the chapter "Das ABC der modernen Logik und Semantik" in:
   Stegmüller, W. (1983) *Erklärung Begründung Kausalität. Probleme und Resultate der Wissenschaftstheorie und Analytischen Philosophie.* Volume I. 2nd Edition. Springer-Verlag, Berlin, Heidelberg, New York.
4. Hempel, C. G. (1965) *Aspects of Scientific Explanation. And Other Assays in the Philosophy of Science.* The Free Press, Collier-Macmillan, New York, London.
5. Hempel, C. G. (1977) *Aspekte wissenschaftlicher Erklärung.* W. de Gruyter, Berlin, New York.
6. Stegmüller, W. (1970) *Erfahrung, Festsetzung, Hypothese und Einfachheit in der wissenschaftlichen Begriffs- und Theorienbildung* [Probleme und Resultate der Wissenschaftstheorie und Analytischen Philosophie, Volume II]. Springer-Verlag, Berlin, Heidelberg, New York, chapter I.
7. Stevens, S. S. (1946) On the theory of scales of measurement. Science *103*, 677–680.
8. Dybkaer, R. & Jörgensen, K. (1989) Measurement, value, and scale. Scand. J. Clin. Lab. Invest. *49*, Supplementum 194, pag. 69–76.
9. Dybkaer, R. (1978) IFCC Approved Recommendation (1978) Quantities and Units in Clinical Chemistry. J. Clin. Chem. Clin. Biochem. *17*, 807–821.
10. Dybkaer, R. (1979) IFCC Approved Recommendation (1978) List of Quantities in Clinical Chemistry. J. Clin. Chem. Clin. Biochem. *17*, 822–835.
11. Büttner, J. (1991) Philosophy of measurement by means of immunoassays. Scand. J. Clin. Lab. Invest. *51*, Supplement 205, 11–20.
12. Beck, J. R. & Pauker, S. G. (1983) The Markov process in medical prognosis. Medical Decision Making [Philadelphia] *2*, 419–458.
13. Myers, L. E., Paulson, D. F., Berry, W. R., Cox, E. B., Laslo, J. & Stanley, W. (1980) A time-dependent statistical model which relates current clinical status to prognosis: application to advanced prostatic cancer. J. Chronic Diseases *33*, 491–499.
14. Peirce, C. S. (1983) *Phänomen und Logik der Zeichen* (Pape, H., ed.) [Suhrkamp Taschenbuch Wissenschaft 425]. Suhrkamp Verlag, Frankfurt.
15. Peirce, C. S. (1986–1990) Semiotische Schriften. Herausgegeben und übersetzt von C. Kloesel und H. Pape. Volume 1 [published 1986], Volume 2 (1903–1906) [published 1990]. Suhrkamp Verlag, Frankfurt.
16. Morris, C. W. (1971) *Writings on the General Theory of Signs.* Mouton, The Hague.
17. Eco, U. (1977) *Zeichen. Einführung in einen Begriff und seine Geschichte.* Suhrkamp Verlag, Frankfurt.
18. Seboek, T. A. (1979) *Theorie und Geschichte der Semiotik* (rowohlts deutsche enzyklopädie). Rowohlt, Reinbek.
19. Seboek, T. A. (General Editor) (1986) *Encyclopedic Dictionary of Semiotics* (Tome 1 to 3). Mouton de Gruyter, Berlin, New York, Amsterdam.
20. The nomenclature used in this paper goes back to *Charles Morris.* See o. c. (16).
21. Baldinger, K. (1957) *Die Semasiologie. Versuch eines Überblicks.* [Deutsche Akademie der Wissenschaften zu Berlin. Vorträge und Schriften, Heft 61]. Akademie-Verlag, Berlin.
22. See o. c. (16) and Morris, C. W. (1979) *Grundlagen der Zeichentheorie. Ästhetik der Zeichentheorie* [Ullstein Materialien 35006] Ullstein, Frankfurt etc., pp. 23 ff.
23. Büttner, H., Hansert, E. & Stamm, D. (1970) Auswertung, Kontrolle und Beurteilung von Meßergebnissen. In: *Methoden der enzymatischen Analyse* (Bergmeyer, H. U., ed.) 2nd Edition. Verlag Chemie, Weinheim. Volume I, pp. 281–364.
24. Stamm, D. & Büttner, J. (1989) Medizinische Beurteilung. In: *Lehrbuch der Klinischen Chemie und Pathobiochemie* (Greiling, H. & Gressner, A. M., eds.) 2nd Edition. Schattauer, Stuttgart, New York, pp. 65–70.
25. Solberg, H. E. (1987) Approved Recommendation (1986) on the theory of reference values. Part 1. The concept of reference values. J. Clin. Chem. Clin. Biochem. *25*, 337–342.
26. Büttner, J. (1990) Leitgedanken in der Geschichte der Klinischen Chemie. Medizinhist. Journal *25*, 268–285.
27. Stettler, A. (1987) Zeichen lesen und Zeichen deuten. Zur Geschichte der Medizischen Semiotik. Gesnerus *44*, 33–54.
28. Leiber, B. & Olbrich, G. (1966) Zur Entwicklungsgeschichte, Definition und Nomenklatur des Syndrombegriffes. In: *Die Klinischen Syndrome* (Leiber, B. & Olbrich, G., eds.) Volume 1: Syndrome. 4th Edition. Urban & Schwarzenberg, München, Berlin, Wien.
29. See mainly o. c. (3c), (4) and (5).
30. Stegmüller, W. (1983) *Erklärung Begründung Kausalität. Probleme und Resultate der Wissenschaftstheorie und Analytischen Philosophie.* Volume I. 2nd Edition. Springer-Verlag, Berlin, Heidelberg, New York.
31. Sadegh-Zadeh, K. (1972) Zur Logik und Methodologie der ärztlichen Urteilsbildung. Meth. Inform. Med. *11*, 203–212.
32. See o. c. (4) and (5).
33. See o. c. (4) and (5).
34. O. c. (30), chapter IX, pp. 774 ff.
35. Carnap, R. (1958) *Induktive Logik und Wahrscheinlichkeit.* Bearbeitet von W. Stegmüller. Springer-Verlag, Wien.
36. For the concept of medical indication see e. g.:
    Anschütz, F. (1982) *Indikation zum ärztlichen Handeln* [Heidelberger Taschenbücher 218]. Springer-Verlag, Berlin, Heidelberg, New York.
37. See l. c. (31).
38. Gross, R. (1973) Analyse des ärztlichen Diagnostikvorganges. In: *Computerunterstützte ärztliche Diagnostik* (Lange, H.-J. & Wagner, G., eds.) Schattauer Verlag, Stuttgart, New York, pp. 31–38.
39. Gross, R. (1973) Einige logische Grundlagen und Grundfragen der Medizin. Deutsches Ärzteblatt *70*, 2319–2321, 2392–2395, 2462–2464, 2538–2540, 2605–2060.
40. Wieland, W. (1975) *Diagnose. Überlegungen zur Medizintheorie.* W. de Gruyter, Berlin, New York.
41. Tautu, P. & Wagner, G. (1978) The process of medical diagnosis: routes of mathematical investigations. Meth. Inform. Med. *17*, 1–10.
42. Gross, R. (1979) Zur Gewinnung von Erkenntnissen in der Medizin. Erfahrung, Intuition, Modelle. Deutsches Ärzteblatt *76*, 2571–2578.
43. Wagner, G., Tautu, P. & Wolber, U. (1978) Problems in medical diagnosis. A bibliography. Meth. Inform. Med. *17*, 55–74.
44. Hartmann, F. (1980) Stellenwert klinisch-chemischer Befunde in verschiedenen Zusammenhängen ärztlicher Urteilsbildung. In: *Validität klinisch-chemischer Befunde* (Lang, H., Rick, W. & Büttner, H., eds.) Merck-Symposium 1979. Springer-Verlag, Berlin, Heidelberg, New York, pp. 8–18.
45. Medawar, P. B. (1969) *Induction and Intuition in Scientific Thought* (Jayne Lectures for 1968). American Philosophical Society, Philadelphia.
46. See also l. c. (38).

47. See o. c. (40), p. 66.
48. See o. c. (40).
49. Büttner, J. (1988) Verwendung klinisch-chemischer Untersuchungen für prognostische Aussagen. In: *Arzneimitteltherapie und Krankheitsprognose.* (Kleinsorge, H. & Schölmerich, P., eds.) G. Fischer Verlag, Stuttgart, New York, pp. 21–43.
50. Pui, C. H., Dodge, R. K., Dahl, G. V., Rivera, G., Look, A. T., Klawinsky, D., Bowman, W. P., Ochs, J., Abromowitch, M., Mirro, J. & Murphy, J. (1985) Serum lactate dehydrogenase level has prognostic value in childhood acute lymphatic leukemia. Blood *66*, 778–782.
51. Büttner, J. (1977) Die Beurteilung des diagnostischen Wertes klinisch-chemischer Untersuchungen. J. Clin. Chem. Clin. Biochem. *15*, 1–12.
52. Büttner, H. (1980) Grundlagen der Bewertung klinisch-chemischer Befunde. In: *Validität klinisch-chemischer Befunde* (Lang, H., Rick, W. & Büttner, H., eds.) Merck-Symposium 1979. Springer, Berlin, Heidelberg, New York, pp. 58–72.
53. Büttner, J. (1989) Die Validität klinisch-chemischer Befunde. In: *Lehrbuch der Klinischen Chemie und Pathobiochemie* (Greiling, H. & Gressner, A. M., eds.) Schattauer, Stuttgart, New York, pp. 71–78.
54. Currie, L. A. (1978) Sources of error and the approach to accuracy in analytical chemistry. In: *Treatise on Analytical Chemistry* (Kolthoff, I. M. & Elving, P. J., eds.) 2nd Edition, J. Wiley & Sons, New York. Part I, vol. 1, pp. 95–242.
55. Büttner, H. (1967) Statistische Qualitätskontrolle in der Klinischen Chemie. Theoretische Grundlagen und praktische Durchführung. Z. Klin. Chem. Klin. Biochem. *5*, 41–48.
56. Büttner, J., Borth, R., Boutwell, J. H., Broughton, P. M. & Bowyer, R. C. (1980) IFCC. Approved Recommendation (1978) on quality control in clinical chemistry. Part 1: General principles and terminology. J. Clin. Chem. Clin. Biochem. *18*, 69–77.
57. L. c. (23).
58. Wisser, H. (1989) Einflußgrößen und Störfaktoren. In: *Lehrbuch der Klinischen Chemie und Pathobiochemie* (Greiling, H. & Gressner, A. M., eds.). 2nd Edition. Schattauer, Stuttgart, New York, pp. 38–57.
59. Büttner, J. (1989) Plausibilitätskontrolle. In: *Lehrbuch der Klinischen Chemie und Pathobiochemie* (Greiling, H. & Gressner, A. M., eds.). 2nd Edition. Schattauer, Stuttgart, New York, pp. 65–68.
60. Porth, A. J., Badke, C., Bothung, S. & Worzyk, M. (1989) Result reports from large centralized laboratories. In: *Data Presentation — Interpretation* (Keller, H. & Trendelenburg, C., eds.) [Clinical Biochemistry, Vol. 2]. W. de Gruyter Verlag, Berlin, New York, pp. 33–61.
61. Guder, W. C. (1980) Einflußgrößen und Störfaktoren bei klinisch-chemischen Untersuchungen. Internist *21*, 533–542.
62. Keller, H. (1980) Einflüsse auf klinisch-chemische Meßgrößen. In: *Validität klinisch-chemischer Befunde* (Lang, H., Rick, W. & Büttner, H., eds.) Merck-Symposium 1979. Springer-Verlag, Berlin, Heidelberg, New York, pp. 25–57.
63. Keller, H., Guder, W. G., Hansert, E. & Stamm, D. (1985) Biological influence factors in clinical chemistry: General considerations. J. Clin. Chem. Clin. Biochem. *23*, 3–6.
64. L. c. (23).
65. L. c. (58).
66. L. c. (23).
67. Gräsbeck, R., Alström, T. & Solberg, H. E. (Editors) (1981) *Reference Values in Laboratory Medicine. The Current State of the Art.* J. Wiley & Sons, Chichester, New York etc.
68. L. c. (51).
69. Miller, M. C., Westphal, M. C. & Reigart, J. R. (1981) *Mathematical Models in Medical Diagnosis.* Praeger Publishers, New York.
70. Griner, P. F., Mayewski, R. J., Mushlin, A. I. & Greenland, P. (1981) Selection and interpretation of diagnostic tests and procedures. Ann. Int. Med. *94*, 553–600.
71. Benson, E. S., Connelly, D. P. & Burke, M. D. (Editors) (1982) *Test Selection Strategies* [Clinics in Laboratory Medicine, vol. 2, no. 4]. W. B. Saunders Co., Philadelphia etc.
72. Abel, U. (1988) *Die statistische Auswertung von Markerdaten in der Onkologie* [Habilitationsschrift, Fakultät für Theoretische Medizin, Universität Heidelberg]. Heidelberg.
73. Extensive bibliography see l. c. (43).
74. Yerushalmy, J. (1947) Statistical problems in assessing methods of medical diagnosis. With special reference to X-ray techniques. Public Health Rep. *62*, 1432–1449.
75. McNicol, D. (1972) *A Primer of Signal Detection Theory.* George Allen & Unwin Ltd., London etc.
76. Swets, J. A. & Pickett, R. M. (1982) *Evaluation of Diagnostic Systems. Methods from Signal Detection Theory.* Academic Press, New York, London etc.
77. Büttner, J. (1982) Grundlagen der Anwendung der Informationstheorie auf qualitative klinisch-chemische Untersuchungen. Anwendung der Informationstheorie auf klinisch-chemische Untersuchungen, I. J. Clin. Chem. Clin. Biochem. *20*, 477–490.
78. Büttner, J. (1989) Information Theoretical Model of a Clinical Chemical Test. In: *Data Presentation — Interpretation* (Keller, H. & Trendelenburg, C., eds.) (Clinical Biochemistry, Vol. 2). W. de Gruyter, Berlin, New York, pp. 219–246.
79. Klar, R. & Reuter, R. (1983) Vergleich zweier unterschiedlicher Ansätze zur Bestimmung von a-posteriori-Kenngrößen für die Bewertung diagnostischer Tests. In: *Methoden der Statistik und Informatik* (Berger, J. & Höhne, K. H., eds.). Springer-Verlag, Berlin a. o., pp. 407–413.

Prof. Dr. Dr. J. Büttner
Institut für Klinische Chemie I
Medizinische Hochschule
Konstanty-Gutschow-Straße 8
W-3000 Hannover 61
Bundesrepublik Deutschland